

RECEIVED

JAN 09 2002

TECH CENTER 1600/2900 1652 1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/246,451A

DATE: 12/20/2001

TIME: 08:56:29

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

ENTERED

C--> 4 <110> APPLICANT: California Institute of Technology;  
 5 Frances H. Arnold  
 6 Hyun Joo  
 8 <120> TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
 11 <130> FILE REFERENCE: 9373/1E827-US1  
 13 <140> CURRENT APPLICATION NUMBER: US 09/246,451A  
 14 <141> CURRENT FILING DATE: 1999-02-09  
 16 <150> PRIOR APPLICATION NUMBER: US 60/094,403  
 17 <151> PRIOR FILING DATE: 1998-07-28  
 19 <150> PRIOR APPLICATION NUMBER: US 60/106,840  
 20 <151> PRIOR FILING DATE: 1998-11-03  
 22 <150> PRIOR APPLICATION NUMBER: US 60/086,206  
 23 <151> PRIOR FILING DATE: 1998-05-21  
 25 <150> PRIOR APPLICATION NUMBER: US 60/106,834  
 26 <151> PRIOR FILING DATE: 1998-11-03  
 28 <160> NUMBER OF SEQ ID NOS: 19  
 30 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
 32 <210> SEQ ID NO: 1  
 33 <211> LENGTH: 1402  
 34 <212> TYPE: DNA  
 35 <213> ORGANISM: P. Putida  
 37 <400> SEQUENCE: 1

38	ctgcaggatc gttatccgct ggccgatctg atcaccacgc gtttttccat cgacgaggcc	60
39	agcaaggcac ttgaactggt caaggcagga gcaactgatc aaccctgat cgactccact	120
40	ctttagccaa cccgcgttcc aggagaacaa caacaatgac gactgaaacc atacaaagca	180
41	acgccaatct tgcccctctg ccaccccatg tgccagagca cctgggtattc gacttcgaca	240
42	tgtacaatcc gtogaatctg tctgcccgcg tgcaggaggc ctgggagatt ctgcaagaat	300
43	caaacgtacc ggatctggtg tggactcgct gcaacggcgg aactggatc gccactcgcg	360
44	gccaactgat ccgtgaggcc tatgaagatt accgccactt ttccagcgag tgcccgttca	420
45	tccctcgtga agccggcgaa gcctacgact tcattcccac ctcgatggat ccgcccgagc	480
46	agcgccagtt tcgtgcgctg gccaaccaag tggttggcat gccgggtggtg gataagctgg	540
47	agaaccggat ccaggagctg gcctgctcgc tgatcgagag cctgcgcccg caaggacagt	600
48	gcaacttcac cgaggactac gccgaaccct tcccgatag catcttcag ctgctcgcag	660
49	gtctaccgga agaagatata ccgcacttga aatacctaac ggaacagatc acccgtccgg	720
50	atggcagcat gaccttcgca gaggccaagg aggcgtctca cgactatctg ataccgatca	780
51	tcgagcaacg caggcagaag ccgggaaccg acgctatcag catcggttgc aacggccagg	840
52	tcaatgggcg accgatcacc agtgacgaag ccaagaggat gtgtggcctg ttactggctg	900
53	gcggcctgga tacggtggtc aatttcctca gcttcagcat ggagttcctg gccaaaagcc	960
54	cggagcatcg ccaggagctg atcgagcgtc ccgagcgtat tccagccgct tgcgaggaac	1020
55	tactccggcg cttctcgtctg gttgcccgtg gccgcactct cacctccgat tacgagtttc	1080
56	atggcgtgca actgaagaaa ggtgaccaga tctgtctacc gcagatgctg tctggcctgg	1140
57	atgagcgcga aaacgcctgc ccgatgcacg tcgacttcag tcgcaaaaag gtttcacaca	1200
58	ccacctttgg ccacggcagc catctgtgcc ttggccagca cctggcccgc cgggaaatca	1260
59	tgtcaccct caaggaatgg ctgaccagga ttcctgactt ctccattgcc ccgggtgccc	1320
60	agattcagca caagagcggc atcgtcagcg gcgtgcaggc actccctctg gtctgggac	1380
61	cggcgactac caaagcggtg ta	1402
63	<210> SEQ ID NO: 2	

Updated  
RSL

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/246,451A

TIME: 08:56:29

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

```

64 <211> LENGTH: 414 ✓
65 <212> TYPE: PRT
66 <213> ORGANISM: P. Putida
68 <400> SEQUENCE: 2
69 Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
70 1 5 10 15
71 His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
72 20 25 30
73 Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
74 35 40 45
75 Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
76 50 55 60
77 Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His
78 65 70 75 80
79 Phe Ser Ser Glu Cys Pro Phe Ile Pro Arg Glu Ala Gly Glu Ala Tyr
80 85 90 95
81 Asp Phe Ile Pro Thr Ser Met Asp Pro Pro Glu Gln Arg Gln Phe Arg
82 100 105 110
83 Ala Leu Ala Asn Gln Val Val Gly Met Pro Val Val Asp Lys Leu Glu
84 115 120 125
85 Asn Arg Ile Gln Glu Leu Ala Cys Ser Leu Ile Glu Ser Leu Arg Pro
86 130 135 140
87 Gln Gly Gln Cys Asn Phe Thr Glu Asp Tyr Ala Glu Pro Phe Pro Ile
88 145 150 155 160
89 Arg Ile Phe Met Leu Leu Ala Gly Leu Pro Glu Glu Asp Ile Pro His
90 165 170 175
91 Leu Lys Tyr Leu Thr Asp Gln Met Thr Arg Pro Asp Gly Ser Met Thr
92 180 185 190
93 Phe Ala Glu Ala Lys Glu Ala Leu Tyr Asp Tyr Leu Ile Pro Ile Ile
94 195 200 205
95 Glu Gln Arg Arg Gln Lys Pro Gly Thr Asp Ala Ile Ser Ile Val Ala
96 210 215 220
97 Asn Gly Gln Val Asn Gly Arg Pro Ile Thr Ser Asp Glu Ala Lys Arg
98 225 230 235 240
99 Met Cys Gly Leu Leu Leu Val Gly Gly Leu Asp Thr Val Val Asn Phe
100 245 250 255
101 Leu Ser Phe Ser Met Glu Phe Leu Ala Lys Ser Pro Glu His Arg Gln
102 260 265 270
103 Glu Leu Ile Glu Arg Pro Glu Arg Ile Pro Ala Ala Cys Glu Glu Leu
104 275 280 285
105 Leu Arg Arg Phe Ser Leu Val Ala Asp Gly Arg Ile Leu Thr Ser Asp
106 290 295 300
107 Tyr Glu Phe His Gly Val Gln Leu Lys Lys Gly Asp Gln Ile Leu Leu
108 305 310 315 320
109 Pro Gln Met Leu Ser Gly Leu Asp Glu Arg Glu Asn Ala Cys Pro Met
110 325 330 335
111 His Val Asp Phe Ser Arg Gln Lys Val Ser His Thr Thr Phe Gly His
112 340 345 350
113 Gly Ser His Leu Cys Leu Gly Gln His Leu Ala Arg Arg Glu Ile Ile

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/246,451A

TIME: 08:56:29

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

```

114          355          360          365
115 Val Thr Leu Lys Glu Trp Leu Thr Arg Ile Pro Asp Phe Ser Ile Ala
116          370          375          380
117 Pro Gly Ala Gln Ile Gln His Lys Ser Gly Ile Val Ser Gly Val Gln
118 385          390          395          400
119 Ala Leu Pro Leu Val Trp Asp Pro Ala Thr Thr Lys Ala Val
120          405          410
122 <210> SEQ ID NO: 3
123 <211> LENGTH: 31
124 <212> TYPE: DNA
125 <213> ORGANISM: Artificial Sequence
127 <220> FEATURE:
128 <223> OTHER INFORMATION: Primer sequence
130 <400> SEQUENCE: 3
131 gatcatgaat gagaccgaca catcacctat c 31
133 <210> SEQ ID NO: 4
134 <211> LENGTH: 32
135 <212> TYPE: DNA
136 <213> ORGANISM: Artificial Sequence
138 <220> FEATURE:
139 <223> OTHER INFORMATION: Primer sequence
141 <400> SEQUENCE: 4
142 acgaattcta gaagaagaaa ctgagggttat tg 32
144 <210> SEQ ID NO: 5
145 <211> LENGTH: 27
146 <212> TYPE: DNA
147 <213> ORGANISM: Artificial Sequence
149 <220> FEATURE:
150 <223> OTHER INFORMATION: Primer sequence
152 <400> SEQUENCE: 5
153 ttggatccgg tggaccttgt ccatttg 27
155 <210> SEQ ID NO: 6
156 <211> LENGTH: 29
157 <212> TYPE: DNA
158 <213> ORGANISM: Artificial Sequence
160 <220> FEATURE:
161 <223> OTHER INFORMATION: Primer sequence
163 <400> SEQUENCE: 6
164 gctctagatc aaccgaagtg cttgtcgag 29
166 <210> SEQ ID NO: 7
167 <211> LENGTH: 24
168 <212> TYPE: DNA
169 <213> ORGANISM: Artificial Sequence
171 <220> FEATURE:
172 <223> OTHER INFORMATION: Primer sequence
174 <400> SEQUENCE: 7
175 cggaattcta ggaaacagac catg 24
177 <210> SEQ ID NO: 8
178 <211> LENGTH: 29

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/246,451A

TIME: 08:56:29

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

```

179 <212> TYPE: DNA
180 <213> ORGANISM: Artificial Sequence
182 <220> FEATURE:
183 <223> OTHER INFORMATION: Primer sequence
185 <400> SEQUENCE: 8
186   ccggatccaa cctgggtcga agtcaaattg
188 <210> SEQ ID NO: 9
189 <211> LENGTH: 24
190 <212> TYPE: DNA
191 <213> ORGANISM: Artificial Sequence
193 <220> FEATURE:
194 <223> OTHER INFORMATION: Primer sequence
196 <400> SEQUENCE: 9
197   catcgatgct taggaggtca tatg
199 <210> SEQ ID NO: 10
200 <211> LENGTH: 25
201 <212> TYPE: DNA
202 <213> ORGANISM: Artificial Sequence
204 <220> FEATURE:
205 <223> OTHER INFORMATION: Primer sequence
207 <400> SEQUENCE: 10
208   tcatgtttga cagcttatca tcgat
210 <210> SEQ ID NO: 11
211 <211> LENGTH: 414
212 <212> TYPE: PRT
213 <213> ORGANISM: Artificial Sequence
215 <220> FEATURE:
216 <223> OTHER INFORMATION: Mutant M7-4H
218 <400> SEQUENCE: 11
219   Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
220     1           5           10           15
221   His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
222           20           25           30
223   Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
224           35           40           45
225   Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
226           50           55           60
227   Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His
228           65           70           75           80
229   Phe Ser Ser Glu Cys Pro Phe Ile Pro Arg Glu Ala Gly Glu Ala Tyr
230           85           90           95
231   Asp Phe Ile Pro Thr Ser Met Asp Pro Pro Glu Gln Arg Gln Phe Arg
232           100          105          110
233   Ala Leu Ala Asn Gln Val Val Gly Met Pro Val Val Asp Lys Leu Glu
234           115          120          125
235   Asn Arg Ile Gln Glu Leu Ala Cys Ser Leu Ile Glu Ser Leu Arg Pro
236           130          135          140
237   Gln Gly Gln Cys Asn Phe Thr Glu Asp Tyr Ala Glu Pro Phe Pro Ile
238           145          150          155          160

```

## RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/246,451A

TIME: 08:56:29

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

```

239 Arg Ile Phe Met Leu Leu Ala Gly Leu Pro Glu Glu Asp Ile Pro His
240           165           170           175
241 Leu Lys Tyr Leu Thr Asp Gln Met Thr Arg Pro Asp Gly Ser Met Thr
242           180           185           190
243 Phe Ala Glu Ala Lys Glu Ala Leu Tyr Asp Tyr Leu Ile Pro Ile Ile
244           195           200           205
245 Glu Gln Arg Arg Gln Lys Pro Gly Thr Asp Ala Ile Ser Ile Val Ala
246           210           215           220
247 Asn Gly Gln Val Asn Gly Arg Pro Ile Thr Ser Asp Glu Ala Lys Arg
248           225           230           235           240
249 Met Cys Gly Leu Leu Val Gly Gly Leu Asp Thr Val Val Asn Phe
250           245           250           255
251 Leu Ser Phe Ser Met Glu Phe Leu Ala Lys Ser Pro Glu His Arg Gln
252           260           265           270
253 Glu Leu Ile Glu Arg Pro Glu Arg Ile Pro Ala Ala Cys Glu Glu Leu
254           275           280           285
255 Leu Arg Arg Phe Ser Leu Val Ala Asp Gly Arg Ile Leu Thr Ser Asp
256           290           295           300
257 Tyr Glu Phe His Gly Val Gln Leu Lys Lys Gly Asp Gln Ile Leu Leu
258           305           310           315           320
259 Pro Gln Met Leu Ser Gly Leu Asp Glu Arg Lys Asn Ala Cys Pro Met
260           325           330           335
261 His Val Asp Phe Ser Arg Gln Lys Val Ser His Thr Thr Phe Gly His
262           340           345           350
263 Gly Ser His Leu Cys Leu Gly Gln His Leu Ala Arg Arg Glu Ile Ile
264           355           360           365
265 Val Thr Leu Lys Glu Trp Leu Thr Arg Ile Pro Asp Phe Ser Ile Ala
266           370           375           380
267 Pro Gly Ala Gln Ile Gln His Lys Ser Gly Ile Val Ser Gly Val Gln
268           385           390           395           400
269 Ala Leu Pro Leu Val Trp Asp Pro Ala Thr Thr Lys Ala Val
270           405           410
272 <210> SEQ ID NO: 12
273 <211> LENGTH: 414
274 <212> TYPE: PRT
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: Mutant M7-6H
280 <400> SEQUENCE: 12
281 Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
282 1           5           10           15
283 His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
284           20           25           30
285 Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
286           35           40           45
287 Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
288           50           55           60
289 Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His
290 65           70           75           80

```

VERIFICATION SUMMARY

DATE: 12/20/2001

PATENT APPLICATION: US/09/246,451A

TIME: 08:56:30

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12192001\I246451A.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date